

IRT observed score による復元得点分布産出プログラムの開発とその応用について

○板宮千尋 (東北大学)
柴山 直 (東北大学)

澁谷拓巳 (東北大学)

キーワード: IRT, Recursion Formula, 学力調査

問題と目的

我が国における大規模学力調査の主たる目的は、学力分布の集団統計量の変化を捉え、(1) 今後の教育施策における見通しを立てること、(2) 児童・生徒の学力向上にむけた指導に役立てることの 2 つである。前者の目的を果たすためには項目を非公開にすることが望ましい一方で、児童・生徒にフィードバックするためには項目の公開が必須であるというジレンマが存在する。このジレンマを解消する 1 つの方法として、項目が公開されるテストとは別に非公開項目からなるテストを実施し、その情報を使って、公開されるテスト間の比較を可能とすることが考えられる。

その際、IRT モデルを利用した等化が有力な方策であるが、日本の大規模学力調査では、受検者に結果を通知する際に素点を用いることが主流となっている (柴山, 2008)。また、1980 年代頃からの学力に関する測定技術の海外における著しい発展にも関わらず、我が国では行政においても研究においても素点をそのまま利用することが主流であり続けてきた (柴山, 2013)。これらの背景から、日本の学力調査に IRT モデルを適用する場合、尺度値 θ ではなく得点を用いて比較することが日本の教育にはより適合すると考えられる。そこで本研究では、IRT モデルにもとづく得点の度数分布を産出するプログラムを開発し、実際に実施された学力調査のデータによる年度間比較を試みた。

方 法

データ 平成 16 年度と平成 18 年度、平成 22 年度に実施された新潟県全県学力調査の数学のデータを利用した。詳細なデータ構造は Figure 1 の通りである。平成 22 年度のデータは、等価グループデザインと係留テストを伴う不等価グループデザインを組み合わせ実施されたため、平成 16 年度と平成 18 年度のデータを共通尺度化する際の共通項目情報として位置づけられる。

各パラメタの推定 IRT モデル (2PLM) における項目パラメタと個人パラメタ (尺度値 θ) の推定には、熊谷 (2017) の EasyEstimation を使用した。なお、今回、尺度値 θ には集団統計量の系統的な推定誤差を避けるために θ_{EAP} から発生させた 10 組の推算値 (Plausible Values) を使用した。

得点分布の産出 IRT モデルを介した得点の算出方法には、ある尺度値 θ の項目反応確率を足し合わせ、期待テスト得点として真値を求める方法と、全ての項目反応パターン考慮する観測得点を求める

方法がある。10 組の推算値を使用した場合と他の尺度値 θ を使用した場合のシミュレーション比較から、受検者数が少ない場合、後者の方が素得点分布に近づき、より正確な分布が得られることが分かった。したがって、Lord 他 (1984) の Recursion Formula に基づく R 言語によるプログラムを開発し、観測得点の分布 (IRT observed score distribution, 復元得点分布; 意識) を産出した。

実施年度	受検者数	版 (α 係数)	項目	
			H16年度問題 (25項目)	H18年度問題 (25項目)
H16	22,804名	1 (.861)		
H18	21,517名	2 (.874)		
H22	568名	1 (.855)		
	559名	2 (.862)		
	568名	3 (.856)		26項目
	568名	4 (.874)		24項目

Figure 1 実施デザイン

結果と考察

平成 16 年度と平成 18 年度の復元得点分布を産出し、比較するために相対累積度数分布を描画したところ、Figure 2 の通りとなった。この結果から、得点の分布が 15 点から 20 点付近でわずかに左寄りに変化しているものの学力分布は全体として大きくは変わらなかったことが分かる。

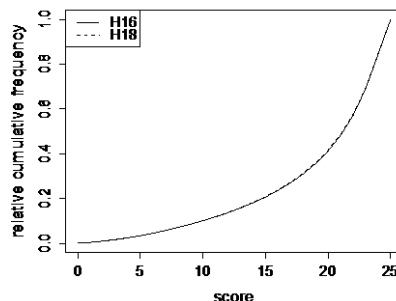


Figure 2 相対累積度数分布

付 記

本研究は JSPS 科研費 16H03731 の助成を受けたものです。また、本研究の遂行にあたり、東北大学教育学部の江尻大亮氏にご協力いただきました。